Intelligent Service Adaptation through AIF Agents

Boris Sedlak, Víctor Casamayor Pujol, Praveen Kumar Donta, Schahram Dustdar

Problem Statement

- ❏ Internet of things (IoT) devices produce **sensory observations**, e.g., video frames in traffic junctions

- ❏ Data **processing** at devices in near vicinity, e.g., visual analysis of traffic composition at junctions

- ❏ **Internal** processing requirements, e.g., response time, that must be continuously evaluated and ensured

- ❏ No **causal understanding** how to ensure requirements; logic confined to distant processing **centers**, i.e., Cloud
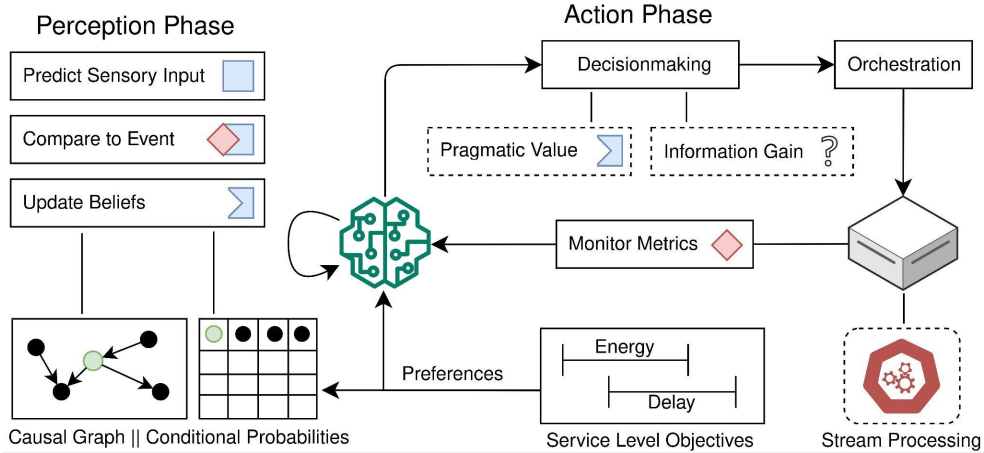
→ Embed AIF agents for decentralized decision-making

Object detection in video frames with Yolov8
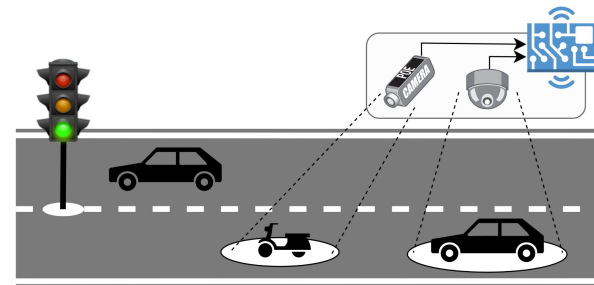
Processing devices are embedded in traffic junctions

Methodology

4. Choose optimal policy to ensure requirements and raise model accuracy

3. Adjust the generative model according to the prediction errors

1. Embed an AIF agent in the processing device for continuous supervision

2. Identify factors that causally impact the requirements fulfillment

Perception Phase

Predict Sensory Input

Compare to Event

Update Beliefs

Causal Graph || Conditional Probabilities

Action Phase

Decisionmaking

Orchestration

Pragmatic Value

Information Gain

Monitor Metrics

Preferences

Energy

Delay

Service Level Objectives

Stream Processing

Outcome

❏ AIF agents **minimize FE** by reconfiguring the processing environment, e.g., change video stream parameter

❏ Individual AIF agents train accurate **generative models** to continuously ensure processing requirements

Pragmatic value (pv)   Information gain (ig)

❏ Models **exchanged** between devices according to their processing capabilities; speeds up device onboarding

❏ Creates **hierarchical structures** that observe and ensure requirements over multiple processing layers